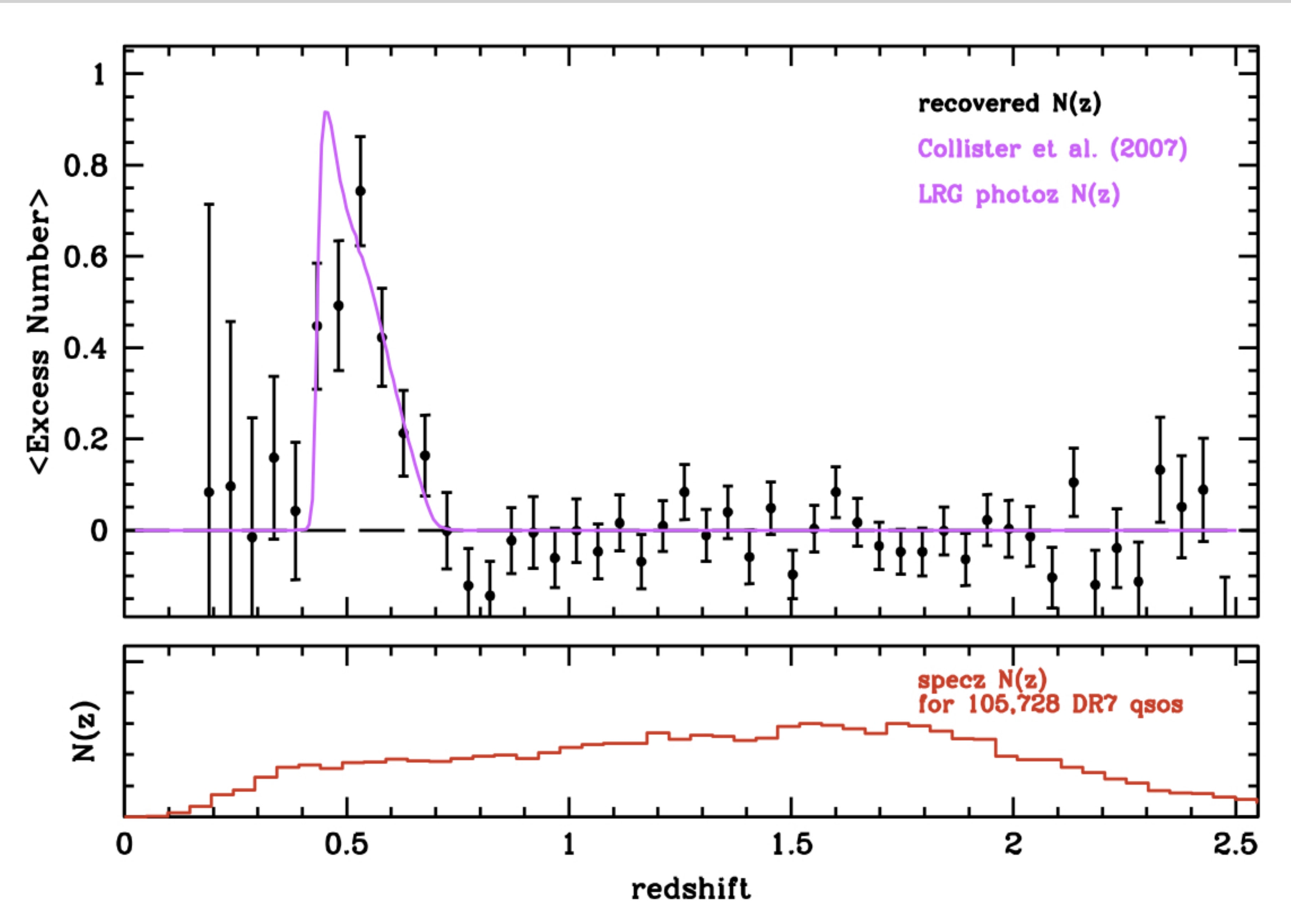
# Inferring Redshift Distributions of Arbitrary Datasets

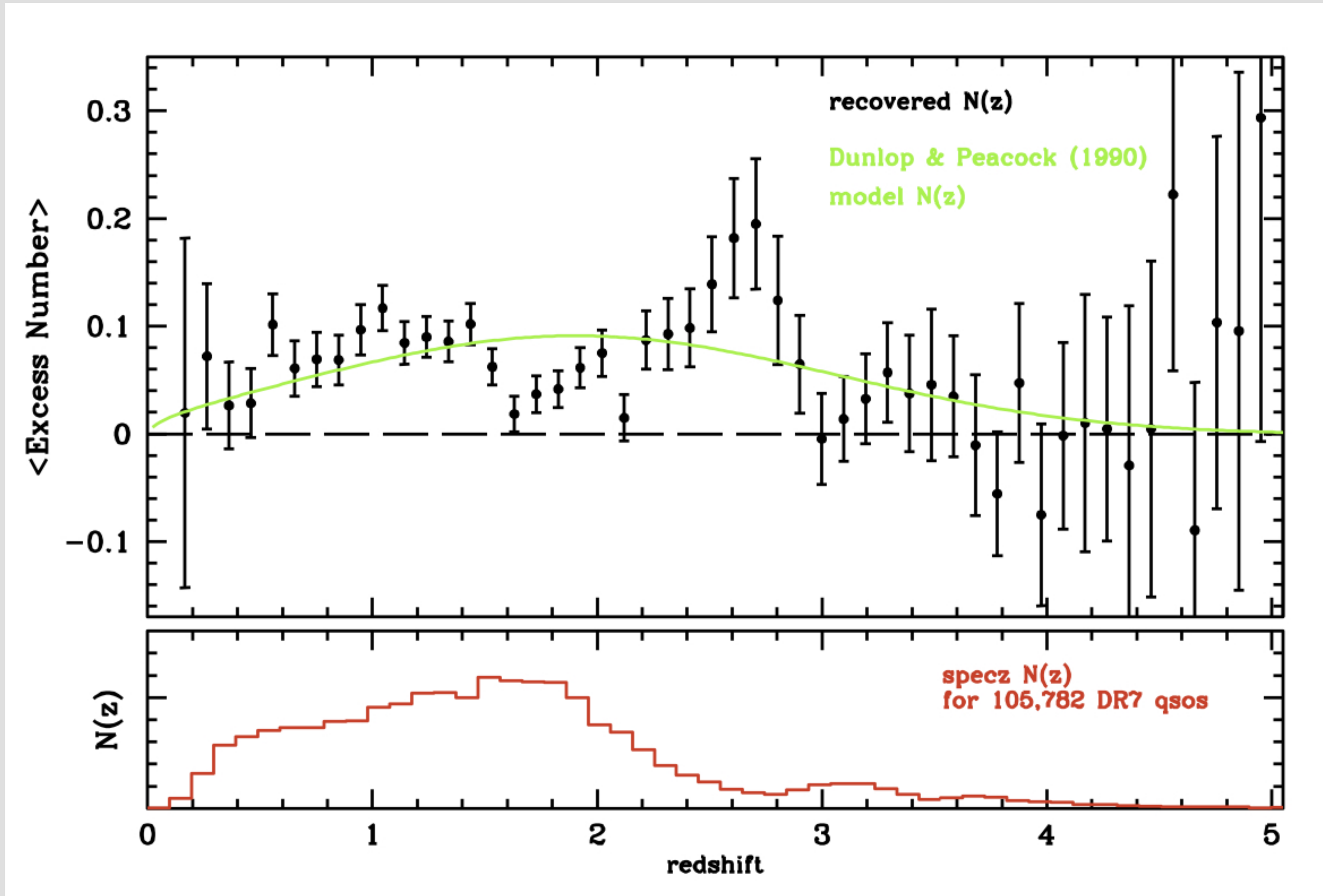
***Samuel Schmidt[1], Brice Menard[2,3], Ryan Scranton[1]***

*[1]University of California, Davis, [2]CITA, University of Toronto, [3]Johns Hopkins University*

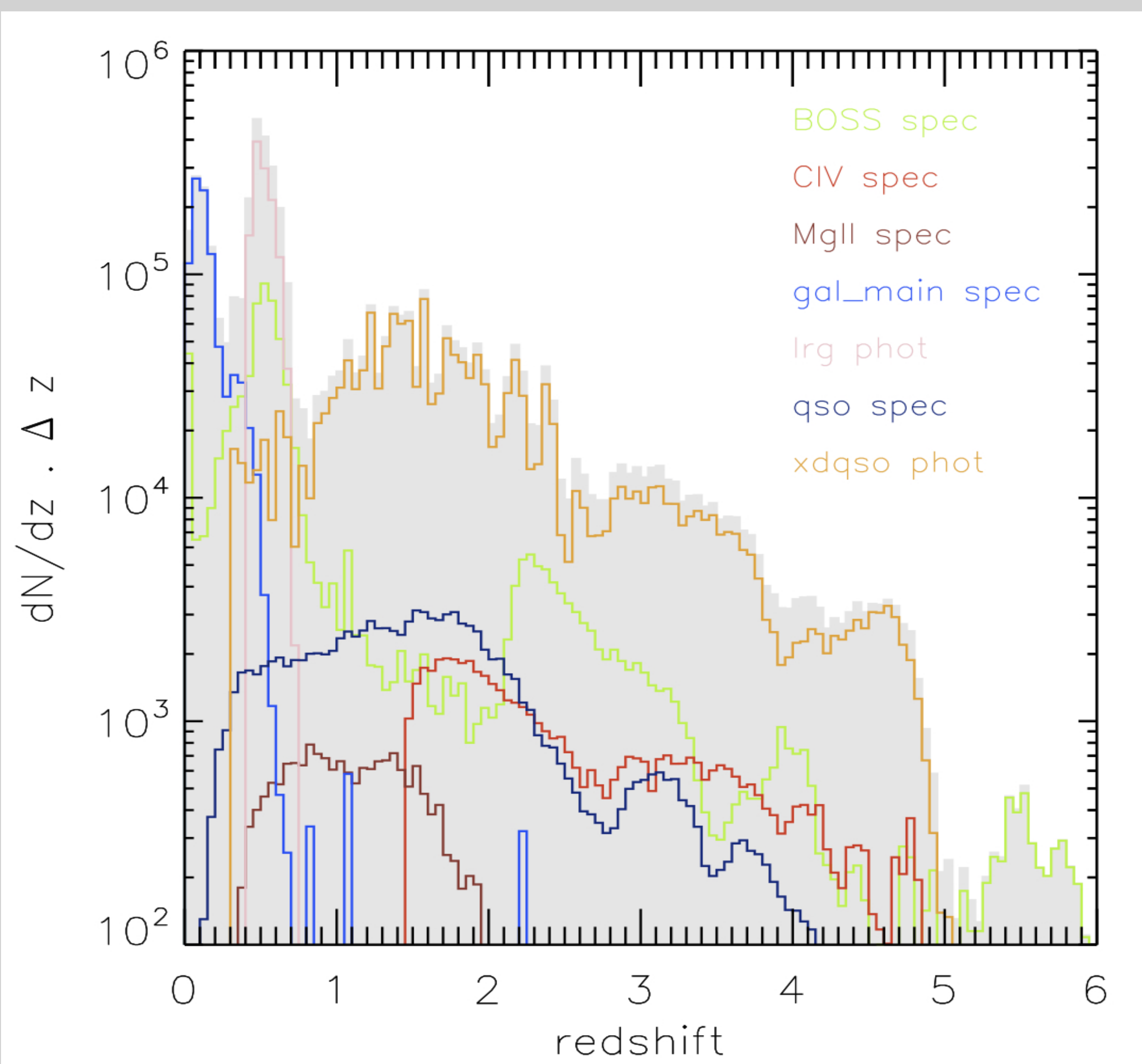
**We present a new tool to measure the redshift distribution of arbitrary astronomical datasets. The method makes use of spatial cross-correlations between an unknown sample and a compilation of existing spectroscopic surveys. It generalizes the standard cross-matching procedure by including spatial information on all scales. This new tool can be applied to both resolved (galaxies) and unresolved (continuous fields & backgrounds) datasets. After describing the method, we present a series of recovered redshift distributions: SDSS photometric quasars, Luminous Red Galaxies, radio sources in the NVSS survey. Finally we will show that this tool is sufficient in scope to estimate redshifts for the upcoming photometric data from LSST and Pan-STARRS.**



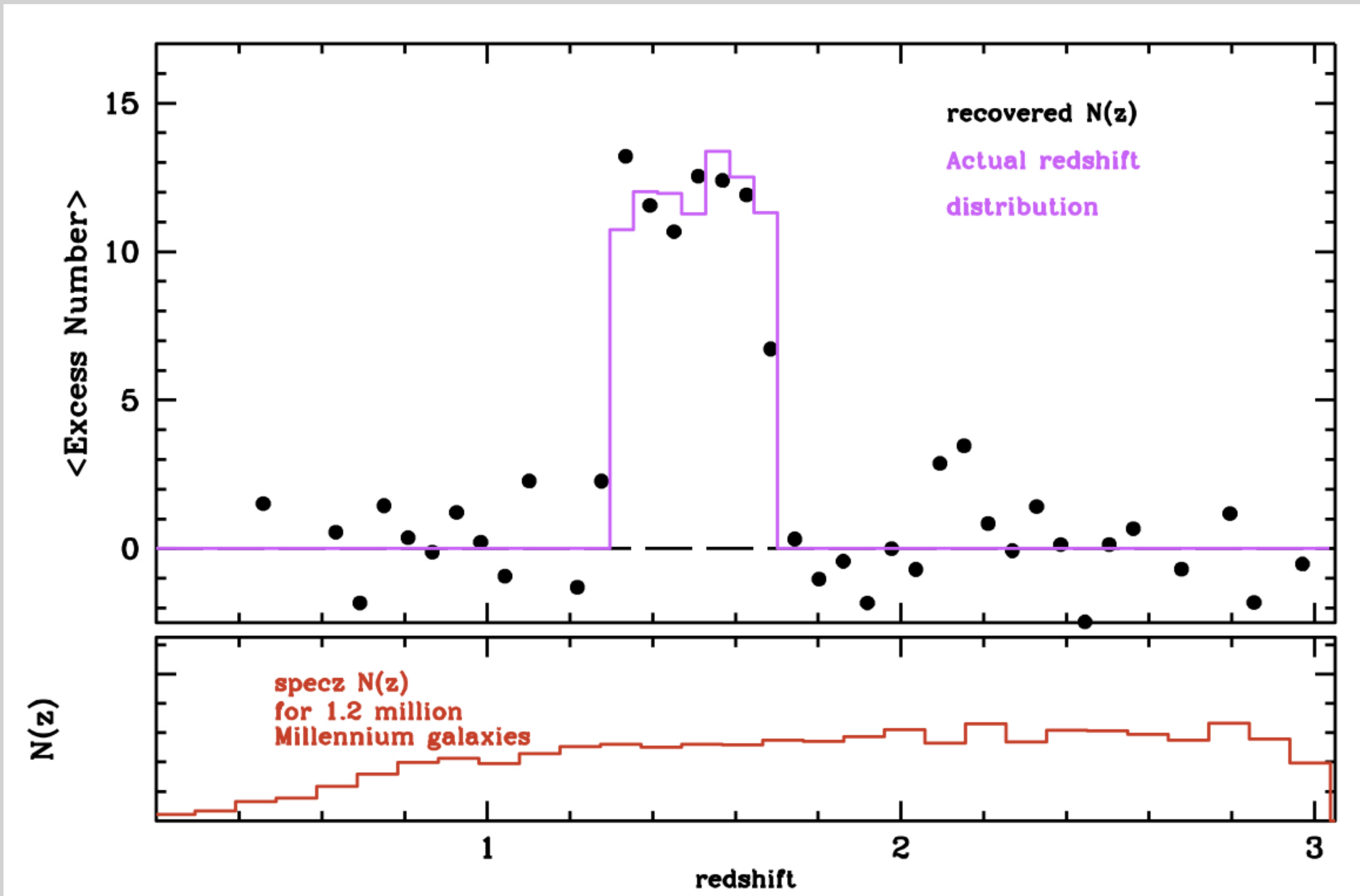

**Reconstruction of the 1.2 million galaxy Mega-Z (Collister et al. 2007) Luminous Red Galaxy (LRG) redshift distribution** formed from cross-correlating against the SDSS DR7 spectroscopic quasar sample (Schneider et al. 2010; red histogram in the lower panel). **The recovered distribution (black) agrees very well with the distribution of photometric redshifts from the same sample (purple curve).** After calibration against a spectroscopic catalog, a photometric redshift sample like the LRGs can be used to bootstrap redshift distributions for other photometric samples with broad and unknown redshift distributions (e.g. the NVSS, shown to the right).




**The reconstruction of the redshift distribution for the NVSS radio survey** (Condon et al. 1998) using SDSS DR7 quasars. The green curve shows the proposed redshift distribution from Dunlop & Peacock (1990) basedon the observed angular clustering of NVSS objects. **Our results suggest that, rather than a single broad distribution, the NVSS objects are actually two nearly distinct samples**, one around z ~ 1 and a higher redshift sample peaking at z ~ 2.7.




**A compilation of spectroscopic and photometric redshift samples that are currently (or soon to be) released.** By including redshifts from quasars (spectroscopic and photometric) as well as Mg II and CIV absorbers and higher redshift LRGs from BOSS, we are able to span a considerably deeper redshift range than the SDSS Main Galaxy sample and at high enough densities that we are able to constrain redshift distributions to z ~ 5. This covers the full expected photometric redshift range for surveys like LSST and Pan-STARRS, meaning that **ambitious faint galaxy-only spectroscopic campaigns will be unnecessary to perform tomographic weak lensing and cluster measurements** for these surveys.




We are currently in the process of **refining our method using Millennium Simulation** light cone (Croton et al. 2006) catalogs. **The plot above shows the reconstruction of a mock tomographic sample from 1.3 < z < 1.7** using a distribution of simulated spectroscopic objects shown in the lower panel, where the two samples were selected to avoid duplicate objects between samples. Cosmic variance is a greater challenge with the simulated data, given its small (here 2x2 degree) footprint. However, by using these samples as a test bed, we will be able to constrain our method's sensitivity to redshift-dependent bias in the spectroscopic sample as well as our assumptions about the underlying cosmology and the comoving scale used.

## References:

Collister et al., 2007, MNRAS, 375, 68
Croton et al, 2006, MNRAS, 365, 11
Newman, 2008, ApJ, 684, 88
Condon et al., 1998, AJ, 115, 1693
Dunlop & Peacock, 1990, MNRAS, 247, 19
Schneider et al, 2010, AJ, 139, 2360